12. Prediction and Decisionmaking in Manual Control*

William R. Ferrell and Harry S. Cohen
Massachusetts Institute of Technology

INTRODUCTION

A simple model of human decisionmaking can be applied to a common class of decision situations, a model sufficiently unrestrictive to have a good chance of representing actual behavior. Among the cases to which the model applies are those classed as signal detection, and a considerable body of experimental evidence substantiates the model's usefulness and its applicability to this and other psychophysical detection tasks. It appears possible to extend the model to include many examples of decisionmaking of practical importance in manual control, those in which an operator seeks to detect not a signal but his own potential success in carrying out an action.

THE MODEL

The decision model and the basic situation to which it applies can be set out very briefly.[1] The situation has the following properties:

(1) There are two mutually exclusive states of nature, S and F, one of which must be the case.

(2) There are observed data d, physical measures, relevant to whether S or F holds.

(3) There are N alternatives, A, B, C . . . , one of which is to be chosen.

(4) The data d generate a scalar p according to a stationary probabilistic process, which is independent of the N alternatives.

(5) The entire p scale is partitioned into N mutually exclusive connected regions each of which is associated with one and only one decision alternative.

(6) When a given observation d' is made, a scalar p' is determined as above and that alternative is chosen which is associated with the region into which p' falls. The p values separating the regions are termed cutoffs.

This simple structure can be related to human decisionmaking by (1) identifying p with the decisionmaker's subjective probability of S given the observation d, and (2) assuming

---

*This work was supported in part by the General Motors Corp. through the General Motor Grant for Highway Safety Research.

[1]Although this statement is new, the model in various forms has been used by many writers. References to sources for the model as well as exposition of ROC curve analysis and its use with human signal detection may be found in D.M. Green, and J.A. Swets, Signal Detection Theory and Psychophysics, John Wiley & Sons, New York, 1966.

that there is a utility associated with each outcome, each combination of chosen alternative and state of nature. If the utilities are known, then by writing expressions for the expected utility of choosing each alternative using the subjective probabilities, cutoffs can be found that will maximize the subjective expected utility. Alternatively, any behavior following the pattern of the model can be interpreted as maximizing expected utility for some utility assignment.

THE RECEIVER OPERATING CHARACTERISTIC CURVE

The performance of a decisionmaker in a given set of circumstances is fully described by the frequencies of the various possible outcomes. When there are two states of nature, S and F, and only two alternatives, A (appropriate to S) and B (appropriate to F), then there are only three independent proportions. The probabilities $Pr(A|S)$, $Pr(A|F)$ and $Pr(S)$ are generally used. When the first two are plotted one against the other, the model generates one point for each cutoff value on the scale, separating the regions associated with the two choice alternatives. The locus of points generated by shifting the cutoff is called the receiver operating characteristic (ROC).

The ROC is extensively used in the analysis of data from psychophysical experiments, especially the detection of signals in noise, for which theoretical curves defining the best possible performance can be obtained. The value of the ROC lies in the fact that it permits a separation of the observer's sensitivity to the true state of nature from the utility considerations affecting his choice. If the ROC curve can be characterized by one parameter then the detector can be given a single index of sensitivity. The operating point along the ROC reflects the decisionmaker's criterion for choosing one alternative over the other and is presumably determined by his utility for the various outcomes.

If the model described above holds, then there are two ways of generating the ROC curve, indirectly through an expression of p the subjective probabilities or directly by recording the outcomes of choices, the rewards for outcomes being varied from time to time to move the cutoff. In the indirect method, the subject observes the data and responds with a number r monotonic with p. For example, he might be asked to estimate his subjective probability of S in percent. From his response and knowledge of whether S or F was the case, the conditional distributions $Pr(r|S)$ and $Pr(r|F)$ are approximated. The experimeter then uses these to infer what would have been the result of decisions made according to the model with various cutoff values, as illustrated in figure 1.

The choice of the numbers to reflect p can, itself, be considered decisionmaking if one can assume that the utility of responding with a number r** when S is the case is greater than that for responding r* if and only if $r^{**} > r^{*}$. If this can be assured by external rewards, no assumption need be made about the direct accessibility of the quantity.

PREDICTING SUCCESS IN MANUAL CONTROL

In an auditory signal detection context an observer listens to a sound that may be either noise alone or noise containing a signal and, on the basis of what he hears, must make a choice between saying "noise" or saying "signal plus noise." The signal's presence or absence has been determined beforehand. The model has been found to be a very good description of human behavior for tasks of this kind. Manual control decisions often show a formal similarity to this paradigm. The operator observes the circumstances, the physical situation, and must decide whether to carry out a particular maneuver, one which he would certainly perform if he know it would be successful. He is thus in the position of trying to detect from a variable set of physical circumstances the success of the action, even though

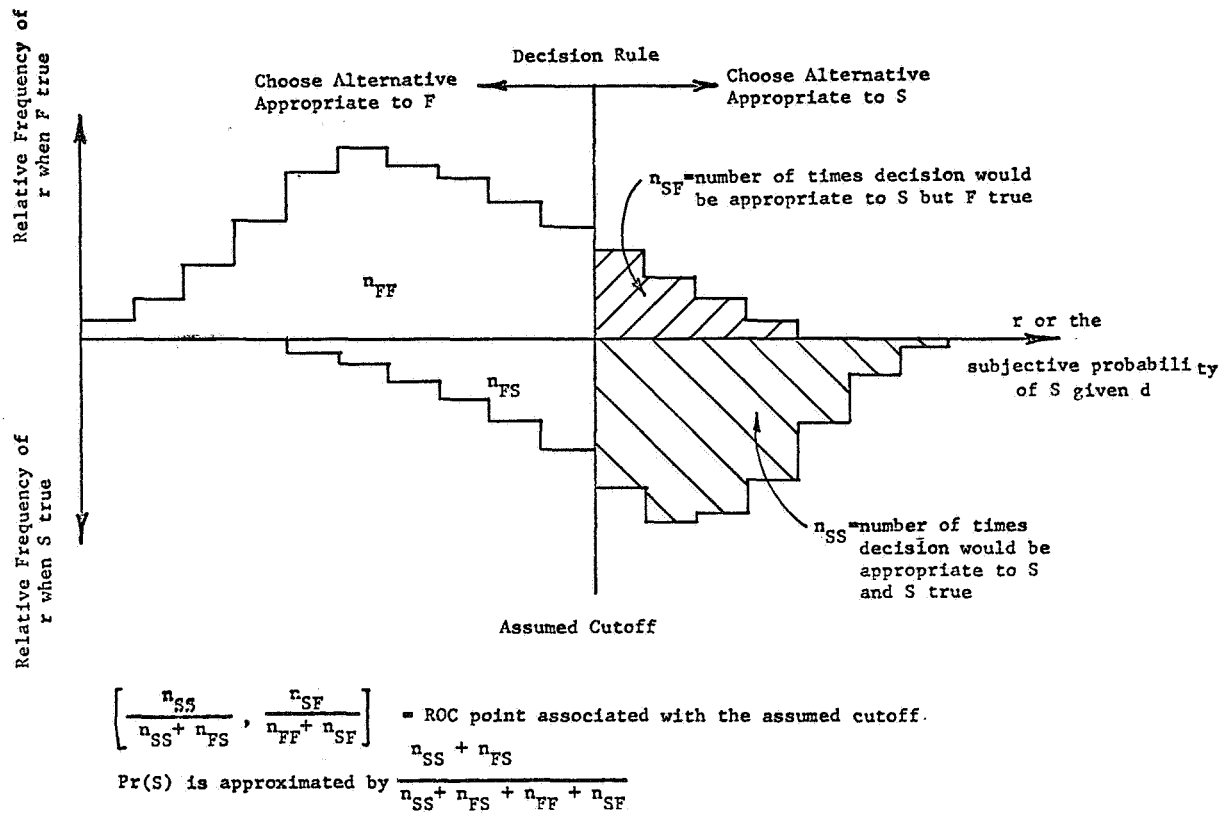
Relative Frequency of r when F true
Decision Rule
Choose Alternative Appropriate to F
Choose Alternative Appropriate to S
n_{SF}=number of times decision would be appropriate to S but F true
n_{FF}
r or the subjective probability of S given d
n_{FS}
Relative Frequency of r when S true
n_{SS}=number of times decision would be appropriate to S and S true
Assumed Cutoff
$\left[\frac{n_{SS}}{n_{SS}+n_{FS}}, \frac{n_{SF}}{n_{FF}+n_{SF}}\right]$ = ROC point associated with the assumed cutoff.
Pr(S) is approximated by $\frac{n_{SS}+n_{FS}}{n_{SS}+n_{FS}+n_{FF}+n_{SF}}$

Figure 1.—Interpretation of data from probability estimation.

whether it will succeed has not yet been determined. Examples are whether to pass the car ahead, whether to continue a landing or make a go around, or whether to try to run an amber light.

In decision situations of this sort, three aspects of performance are relevant, (1) the decisionmaker's skill, (2) his ability to predict the result of his maneuvers, and (3) the relative importance to him of success and failure. If the given model holds, these components can be separately assessed for manual control decisions for which an ROC curve can be determined. Skill is represented by the probability of success, Pr(S). The ability to predict is given by the ROC curve. The operator's effective utility for success relative to that for failure can be found from the operating point on the ROC and Pr(S). If the decisionmaker operates according to the model with a cutoff p_o, his expected objective utility is

$$U_{SS}\ Pr(p \geq p_o | S) + U_{SF}\ Pr(p \geq p_o | F)\ [1-Pr(S)]$$
$$+ U_{FS}\ Pr(p < p_o | S)\ Pr(S) + U_{FF}\ Pr(p < p_o | F)\ [1-Pr(S)]$$

where p is his subjective probability and the utilities are defined by

	S occurs	F occurs
Response consistent with S	U_{SS}	U_{SF}
Response consistent with F	U_{FS}	U_{FF}

The quantities $Pr(p \geq p_o|S)$ and $Pr(p \geq p_o|F)$ are the ordinate and abscissa of the ROC. Rewriting the equation gives curves of constant utility as straight lines with slope equal to

$$\left[\frac{1-Pr(S)}{Pr(S)}\right]\left[\frac{U_{FF}-U_{SF}}{U_{SS}-U_{FS}}\right]$$

If the subjective probability of S equals the objective probability of S, the second term is a measure of the utility for failure relative to that for success. It can be evaluated from Pr(S) and the tangent to the ROC at the operating point. Thus the three components of the decisionmaker's performance can be separately evaluated if the ROC analysis is valid for manual control decisions.

THE EXPERIMENTAL QUESTION

The usefulness of the ROC description of decision behavior in manual control tasks depends on whether real decisionmakers are adequately represented by the decision model. The model predicts that a decisionmaker has an ROC curve that is independent of the particular set of choice alternatives. If points on a decisionmaker's ROC graph generated by various sets of decision alternatives do not lie on the same ROC curve, his decision-making ability cannot be separated from the reward structure of the decision situation.

The existence of stable ROC curves for predicting success in manual control tasks was investigated by comparing points on an ROC graph generated by bet choices with points generated by probability estimation.

EXPERIMENT I

The experimental manual control task consisted of turning a knob to deflect an x-y recorder pen in the y direction past obstacles and into a target area as it moved at constant rate in the x direction. The dynamics were of the form k/S^2. The obstacles and target varied randomly in both size and position. A typical configuration is shown in figure 2. The conditions were—

(1) controller choosing bets
(2) observer choosing bets
(3) controller making probability estimates
(4) observer making probability estimates

On each trial, subjects were first shown the particular obstacles-target configuration. In the bet choice sessions, subjects chose one of two bets, A or B, the payoff of which depended on whether the target was hit. On each trial one or another of the following payoff matrices was used:

	Success	Failure		Success	Failure
Bet A	9¢	0¢	Bet A	7¢	0¢
Bet B	5¢	1¢	Bet B	5¢	3¢

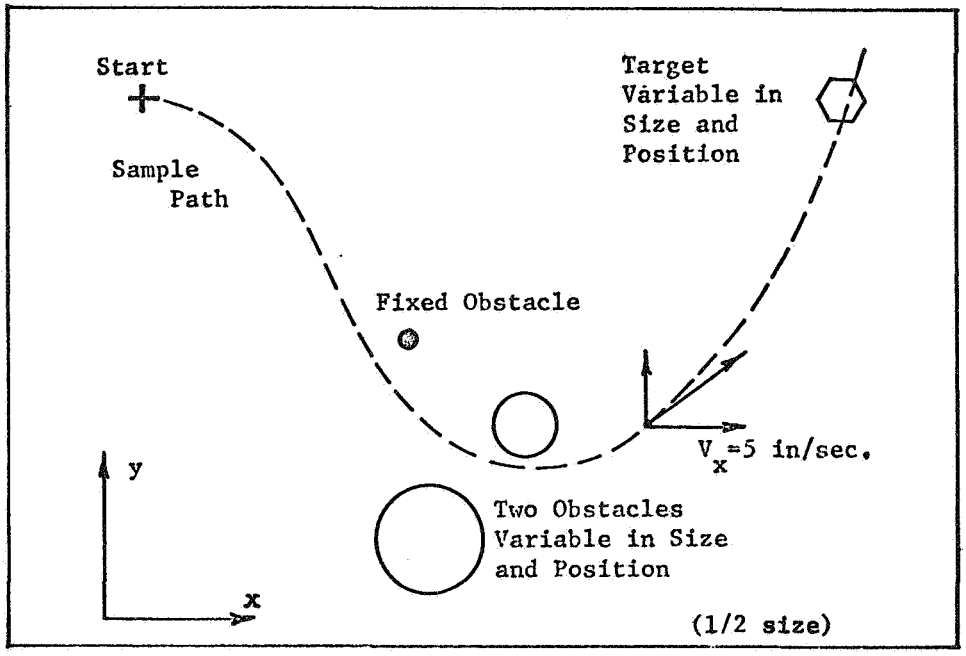

Figure 2.—Task layout for experiment I.

In no case did it pay to fail. In the probability estimation sessions, the subjects picked a number between 0 and 100 representing their feelings about whether or not the task would be performed successfully. Subjects knew that their pay would be determined by decisions chosen optimally using their estimates. Controllers and observers were paired. The controllers each performed the given task, after choosing or estimating, to determine the outcome. The observer paired with the controller was presented task configurations in the same order but saw the outcome by observing the pen trace produced by the controller.

The experimental results are plotted in ROC form in the four graphs of figure 3.

EXPERIMENT II

The manual control task consisted of driving an automobile between three sets of rubber traffic cones on a parking lot. The spacing and offset of the second two pairs of cones varied randomly from trial to trial. A typical arrangement is shown in figure 4. The driver had to enter the first gate at 20 mph, and decisions or estimates had to be made by this time. Knocking over one or more cones was considered a failure.

There were two subjects in this experiment. On each trial, one served as the controller and the other sat beside him in the front seat as an observer. The controller told his decisions to an experimenter in the rear seat and the observer wrote his down. The observer was prevented from hearing the controller's decisions by earphones presenting white noise.

On half the trials, the controller and observer each chose one of two bets defined by the payoff matrix:

	Success	Failure
Bet A	15¢	-10¢
Bet B	5¢	-5¢

On the remaining trials they each chose a number between 0 and 100 to represent their feelings about the chances of success.

The experimental results are plotted in ROC form in figure 5.

EXPERIMENTAL CONCLUSIONS

In both experiments, the points generated by bet choices fell close to the ROC curve generated by probability estimates indicating that for these tasks, the decisionmaker's predictive ability can be meaningfully described by a single ROC curve.

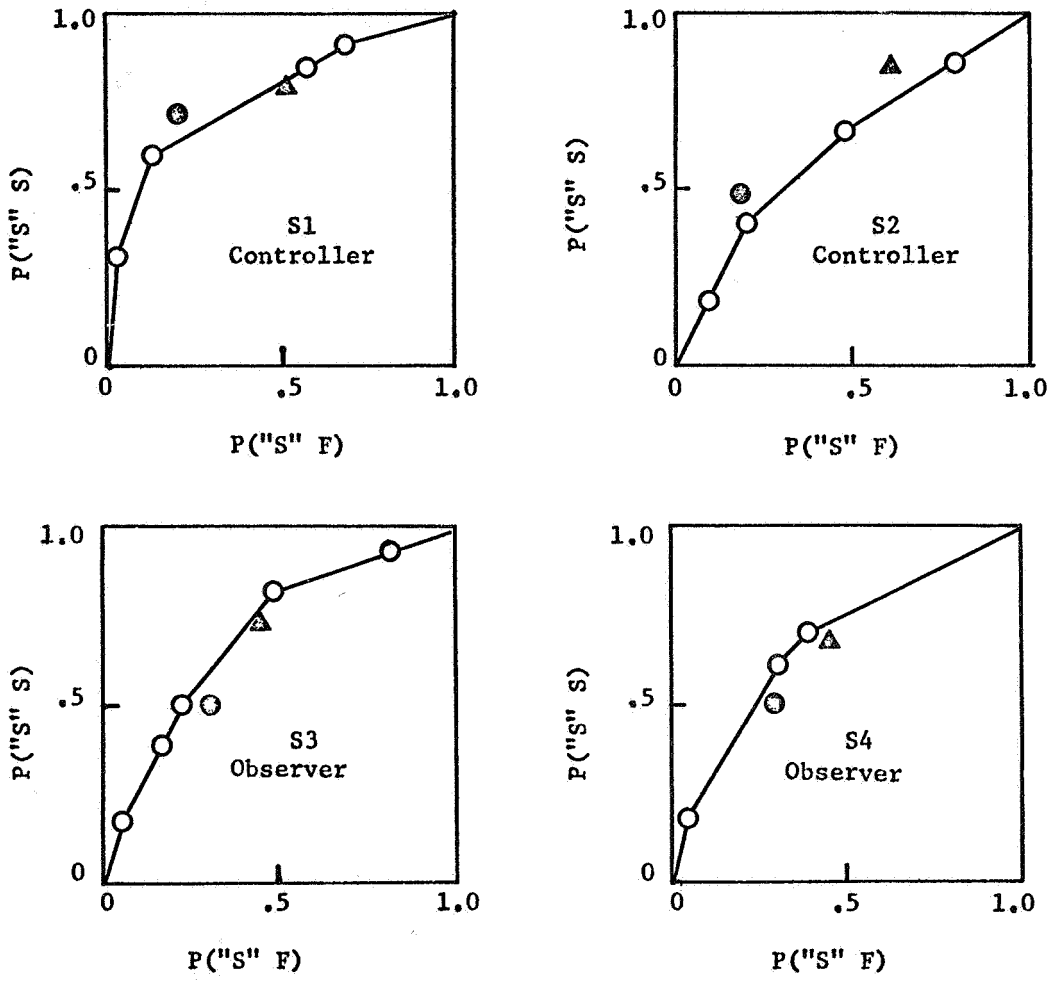

Figure 3.—Results from experiment I.

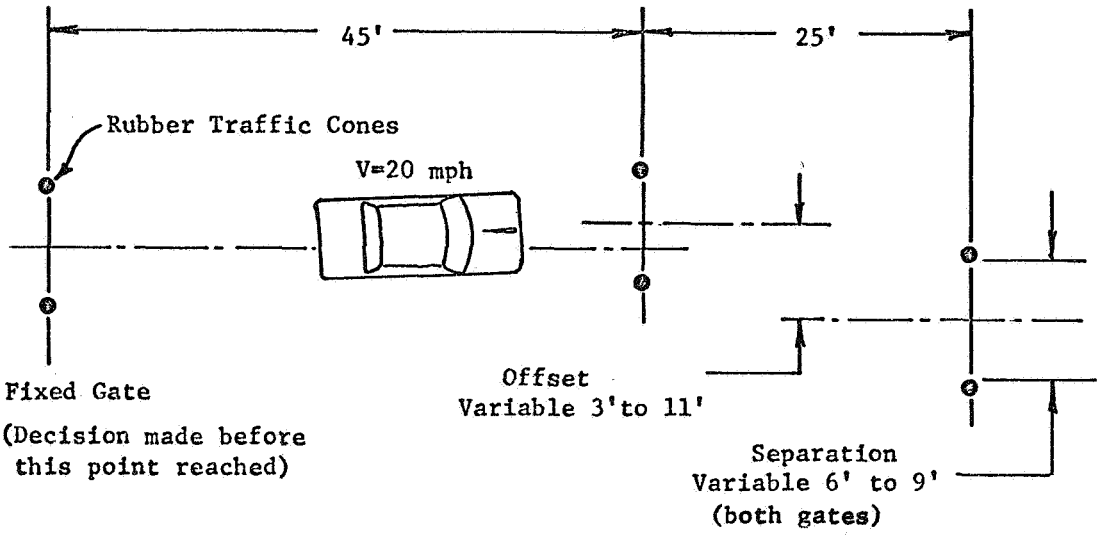

Figure 4.—Task layout for experiment II.

CONTROLLER VERSUS OBSERVER

No significant difference in the shape of the ROC curves of controllers and observers was found. This result indicates that the controller has no further source of information beyond the task configuration and his past performance on which to predict his success. Controllers however tended to locate their operating points above and to the right of observers on their ROC curve.

In terms of the decision model, this means that, though equally able to detect future success, controllers make the choice consistent with success more often. This is much as if they had in addition to monetary payoff an extra, intangible reward for correctly predicting success.

POTENTIAL APPLICATION

VALIDATION OF SIMULATORS.—A simulator is an artificial environment in which events occur analagous to events in the real world. Its effectiveness lies in the strength of the inferences that can be drawn about an operator's behavior in the real world from his behavior in the simulator.

An operator's decisionmaking performance in a simulator will differ from that in an actual situation if either he perceives the causal relations among stimuli differently in the two environments or he has in the two cases entirely different utilities for real and simulated outcomes. The ROC formulation makes it possible to specify precisely the requirements for simulator validity. With respect to the first problem, if the ROC curve for a given task is the same in both environments, then decisionmaking in the simulator for that task is potentially the same as that in the actual situation; for the ROC represents the limiting ability of the decisionmaker to extract relevant information from the data available.

The argument is widely advanced that results from simulators can never be extrapolated with confidence since the operator never risks his life. If it has been established or can be assumed that the same ROC curve holds both in the simulated and real situations, then the effect of a difference in utility for the outcomes under the respective circumstances

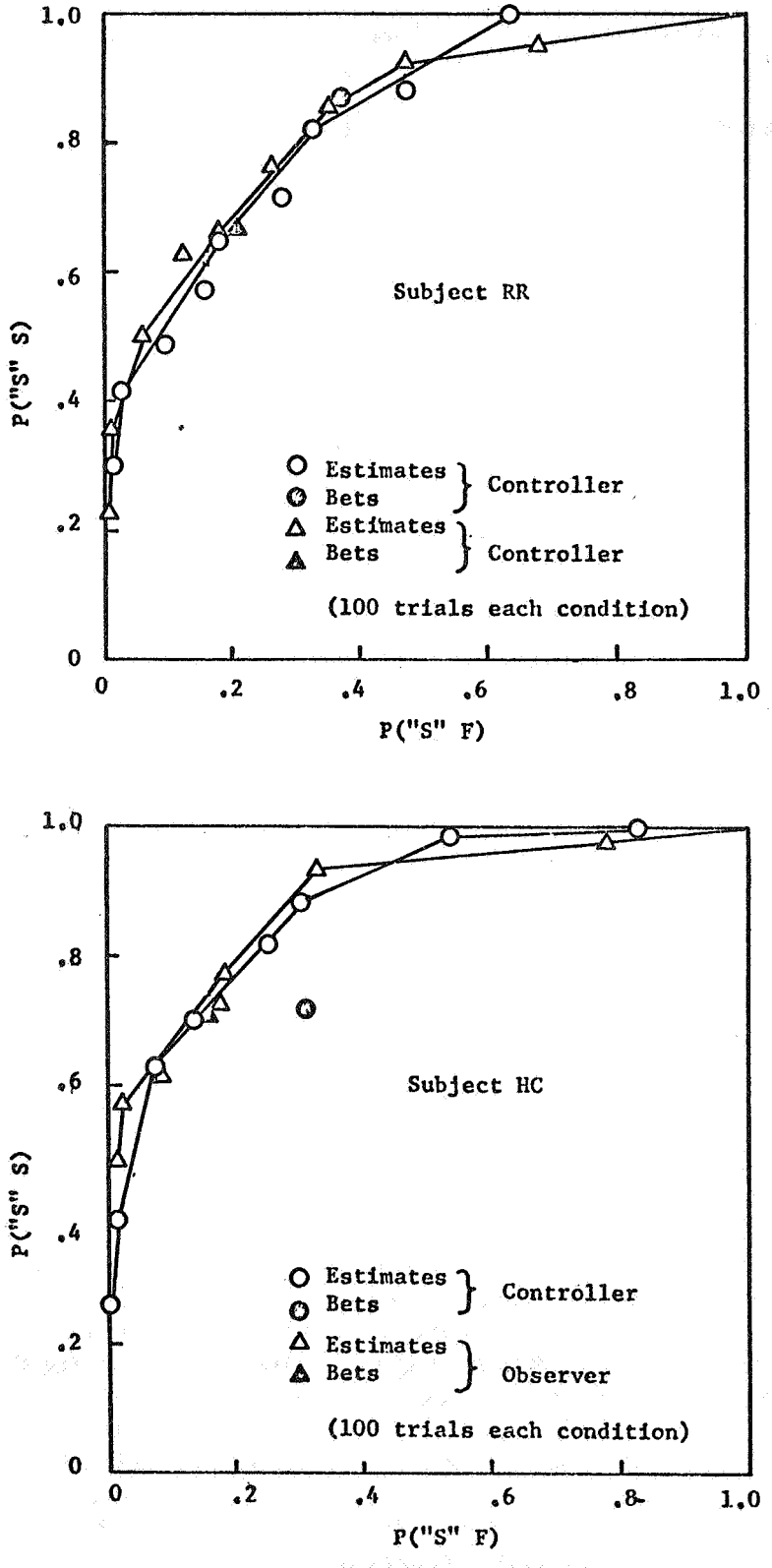

Figure 5.—Results from experiment II.

is solely in the operating point. The greater cost of real failures will result in a more conservative cutoff and the range of operating conditions will be precisely delineated.

EVALUATION OF DISPLAYS. — Man-machine systems sometimes contain displays that augment the man's ability to predict his performance at a control task without influencing his ability to perform the task. These help the man decide whether or not to try a given task. An example is a device that tells a driver how much time remains in the amber phase of a traffic light cycle. This information could help the driver decide whether to stop or try to get through the intersection before the light changes.

The ROC curve is an appropriate measure of predictive ability to be used in the experimental evaluation of such displays, and where necessary enables the separation of the display's enhancement of the operator's control skill from its effect on his predictive ability.